SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

AD-A225 793

DTIC FILE COPY

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AIM 1187	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Direct Recovery of Motion and Shape in the General Case by Fixation		5. TYPE OF REPORT & PERIOD COVERED Memorandum
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) M. Ali Taalebinezhaad		8. CONTRACT OR GRANT NUMBER(s) N00014-85-K-0124
9. PERFORMING ORGANIZATION NAME AND ADDRESS Artificial Intelligence Laboratory 545 Technology Square Cambridge, MA 02139		10. PROGRAM ELEMENT PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Advanced Research Projects Agency 1400 Wilson Blvd. Arlington, VA 22209		12. REPORT DATE March 1990
		13. NUMBER OF PAGES 21
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) Office of Naval Research Information Systems Arlington, VA 22217		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Distribution is unlimited

DTIC ELECTE AUG 23 1990 D

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

Co

18. SUPPLEMENTARY NOTES

None

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

direct motion vision	fixation
structure and motiton	least squares

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Abstract: In motion vision, the problem is to find, from a sequence of time varying images, the relative rotational and translational velocities between a viewer and an environment as well as the shape of objects in the environment. This paper introduces a *direct method* called *fixation* for solving the general motion vision problem. This fixation method results in a constraint equation between

(continued on back)

Block 20 continued:

translational and rotational velocities that in combination with the *Brightness-Change Constraint Equation* solves the general motion vision problem, arbitrary motion with respect to an arbitrary rigid environment.

Avoiding *correspondence* and *optical flow* has been the motivation behind the *direct methods* because both solving the *correspondence* problem, and computing *the optical flow* reliably, have proven to be rather difficult and computationally expensive.

Recently direct motion vision methods, which directly use the image brightness information such as temporal and spatial brightness gradients directly, have used the *Brightness-Change Constraint Equation* for solving the motion vision problem in special cases such as *Known Depth*, *Pure Translation* or *Known Rotation*, *pure Rotation*, *Planar World* and *Quadratic Patches.* In contrast to these solutions, our fixation method does not put such severe restrictions on the motion or the environment.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
ARTIFICIAL INTELLIGENCE LABORATORY

A.I. Memo No. 1187 March 1990

Direct Recovery of Motion and Shape in the General Case by Fixation

M. Ali Taalebinezhaad

Abstract: In motion vision, the problem is to find, from a sequence of time varying images, the relative rotational and translational velocities between a viewer and an environment as well as the shape of objects in the environment. This paper introduces a *direct method* called *fixation* for solving the general motion vision problem. This fixation method results in a constraint equation between translational and rotational velocities that in combination with the *Brightness-Change Constraint Equation* solves the general motion vision problem, arbitrary motion with respect to an arbitrary rigid environment.

Avoiding *correspondence* and *optical flow* has been the motivation behind the *direct methods* because both solving the *correspondence* problem and computing *the optical flow* reliably have proven to be rather difficult and computationally expensive.

Recently direct motion vision methods, which directly use the image brightness information such as temporal and spatial brightness gradients directly, have used the *Brightness-Change Constraint Equation* for solving the motion vision problem in special cases such as *Known Depth*, *Pure Translation* or *Known Rotation*, *pure Rotation*, *Planar World* and *Quadratic Patches*. In contrast to these solutions, our fixation method does not put such severe restrictions on the motion or the environment.

Keywords: Direct Motion Vision, Fixation, Structure and Motion, Least Squares.

Acknowledgments: This paper describes research done at the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology. Support for the laboratory's artificial intelligence research is provided in part by the Advanced Research Projects Agency of the Department of Defense under Office of Naval Research contract N00014-85-K-0124.

1 INTRODUCTION

In motion vision, the goal is to recover, from time varying images, the relative motion between a viewer and an environment as well as the structure of objects in the environment. A survey of previous literatures on machine vision is given by Barron [10]. Some of the current issues in image flow theory and motion vision are discussed by Waxman & Wohn [32] and Aloimonos & Shulman [4]. Much of the earlier work on recovering motion has been based on either establishing correspondences between the images of prominent features (points, lines, contours, and so on) in an image sequence (for example, Prazdny [27], Ullman [30, 31], Longuet-Higgins [19], and Aloimonos & Basu [2]) or establishing the velocity of points over the whole image, commonly referred to as the optical flow (for example, Ballard & Kimball [5], Bruss & Horn [11], and Adiv [1]).

In general, identifying features here means determining gray-level corners. For images of smooth objects, it is difficult to find good features or corners. Further, the *correspondence* problem has to be solved, that is, feature points from consecutive frames have to be matched.

The computation of the local flow field exploits a constraint equation between the local brightness changes and the two components of the *optical flow*. This only gives the components of flow in the direction of the brightness gradient. To compute the full flow field, one needs additional constraints such as the heuristic assumption that the flow field is locally smooth (Hildreth [14], and Horn & Schunck [15]). This leads to an estimated optical flow field that is not the same as the true motion field.

Both solving the *correspondence* problem, and computing *optical flow* reliably, have proven to be rather difficult and computationally expensive. This has motivated the investigation of *direct methods* which use the image brightness information directly to recover motion.

Recently direct motion vision methods have used the *Brightness-Change Constraint Equation* (BCCE) for solving the motion vision problem in special cases such as *Known Depth* [15], *Pure Translation* or *Known Rotation* [17, 24, 16], *Pure Rotation* [17], *Planar World* [23] and *Quadratic Patches* [22]. In this work, a method called *fixation* has been introduced which in combination with the brightness-change constraint equation solves the direct motion vision problem of arbitrary motion with respect to an arbitrary rigid environment. That is, it recovers the shape, rotational velocity and translational velocity

in the general case. In contrast to the tracking methods presented in [3, 9, 28, 29], our *fixation* method is not only different but also is general. For example, Aloimonos & Tsakiris [3] propose a method for tracking a target of known shape. Badopadhay, Chandra & Ballard [9] use optical flow for tracking. Sadini & Tistarelli [29] do tracking for the special case in which the component of rotational velocity along the optical axis is zero.

A block diagram of the ideas behind this work is shown in figure 1. We

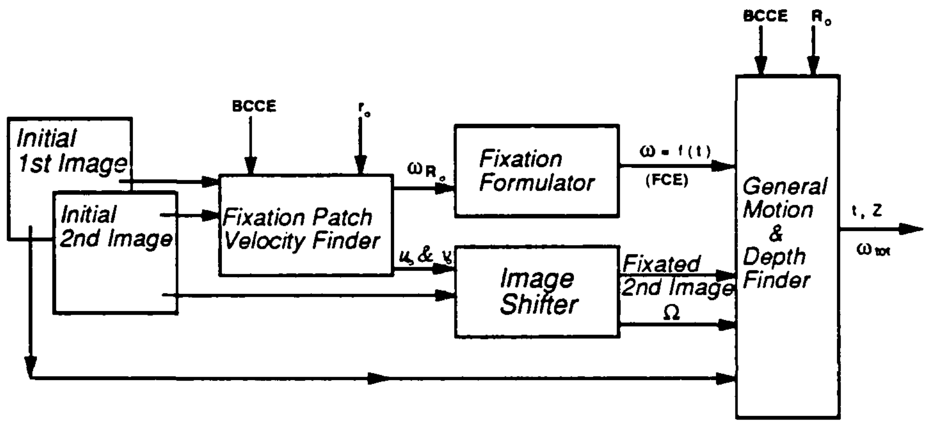

Figure 1: A block diagram of the *fixation method* modules.

start with a brief review of the BCCE in section 2. Then in section 3, we show that by choosing an *interest point* in the environment, $\mathbf{R}_o$, and knowing the component of rotational velocity along the position vector associated with that interest point, $\omega_{\mathbf{R}_o}$, we can obtain a *Fixation Constraint Equation* (FCE) between the *rotational velocity* $\boldsymbol{\omega}$ and the *translational velocity* $\mathbf{t}$ just by keeping the image of the interest point stationary in the image plane. Section 4 shows how the *fixation constraint equation* can be combined with the BCCE and applied to *fixated images* in order to find $\mathbf{t}$, $\boldsymbol{\omega}$ and depth Z in the general case. Recovering $\omega_{\mathbf{R}_o}$, needed in the fixation constraint equation, and finding the components of *fixation velocity*, u_o and v_o, necessary for obtaining a fixated 2nd image, are discussed in section 5. In order to apply the fixation constraint equation, a sequence of two fixated images is needed. Initial 1st image can be used directly and section 6 shows how a fixated 2nd image is obtained from the initial 2nd image.

2 THE BRIGHTNESS CHANGE CONSTRAINT EQUATION

Using a viewer-based coordinate system which is adopted from Longuet-Higgins and Prazdny [18] is very common in direct motion vision. Figure 2 depicts the coordinate system under consideration.

In this coordinate system, a world point

$$\mathbf{R} = (X\ Y\ Z)^T \tag{1}$$

is imaged at

$$\mathbf{r} = (x\ y\ 1)^T. \tag{2}$$

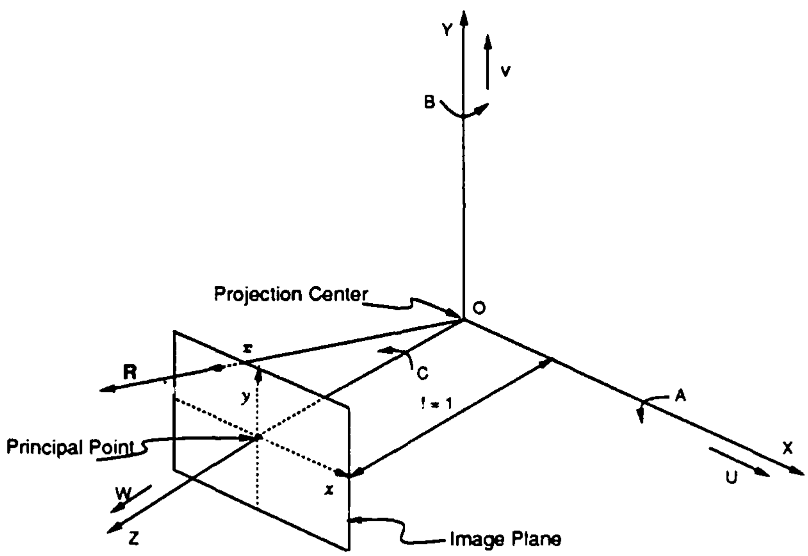

Figure 2: The viewer-centered coordinate system. The translational velocity of the camera is $\mathbf{t} = (U\ V\ W)^T$, and its rotational velocity is $\boldsymbol{\omega} = (A\ B\ C)^T$.

That is, the image plane has the equation $Z = 1$ or in other words the *focal length* is 1. The origin is at the projection center and the Z-axis runs along the optical axis. The X- and Y- axes are parallel to the x- and y- axes of

the image plane. Image coordinates are measured relative to the *principal point*, the point $(0\ 0\ 1)^T$ where the optical axis pierces the image plane. The position vectors $\mathbf{r}$ and $\mathbf{R}$ are related by the perspective projection equation

$$\mathbf{r} = (x\ y\ 1)^T = \left(\frac{X}{Z}\ \frac{Y}{Z}\ \frac{Z}{Z}\right)^T = \frac{\mathbf{R}}{\mathbf{R} \cdot \hat{z}} \tag{3}$$

where $\mathbf{R} \cdot \hat{\mathbf{z}} = Z$ and $\hat{\mathbf{z}}$ denotes the unit vector in the Z direction.

When the observer moves with instantaneous translational velocity $\mathbf{t} = (U\ V\ W)^T$ and instantaneous rotational velocity $\boldsymbol{\omega} = (A\ B\ C)^T$ relative to a rigid environment, then the time derivative of the vector $\mathbf{R}$ can be written as

$$\mathbf{R}_t = -\mathbf{t} - \boldsymbol{\omega} \times \mathbf{R}. \tag{4}$$

The motion of the world point $\mathbf{R}$ results in motion of the corresponding image point $\mathbf{r}$. It can be shown [23, 21] that the motion field in the image plane is obtained by differentiating eqn. (3) with respect to time as

$$\mathbf{r}_t = \frac{d}{dt}\left(\frac{\mathbf{R}}{\mathbf{R} \cdot \hat{\mathbf{z}}}\right) = \frac{\hat{\mathbf{z}} \times (\mathbf{R}_t \times \mathbf{r})}{\mathbf{R} \cdot \hat{\mathbf{z}}}. \tag{5}$$

Substituting for $\mathbf{R}$, $\mathbf{r}$ and $\mathbf{R}_t$ from equations (1), (2) and (4) into eqn. (5) gives, [18, 11]

$$\mathbf{r}_t = \begin{pmatrix} x_t \\ y_t \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{-U+xW}{Z} + Axy - B(x^2+1) + Cy \\ \frac{-V+yW}{Z} - Bxy + A(y^2+1) - Cx \\ 0 \end{pmatrix}. \tag{6}$$

This result is just the *parallax equations of photogrammetry* that occur in the incremental adjustment of relative orientation [13, 20]. It shows how, given the environment motion, the motion field can be calculated for every image point.

Image brightness changes are primarily due to the relative motion between an environment and an observer provided that the surfaces of the objects have sufficient texture and the lighting condition varies slowly enough both spatially and with time. In this case, brightness changes due to changing surface orientation and changing illumination can be neglected. Consequently, we may assume that the brightness of a small patch on a surface in the scene

does not change during motion. Then expansion of the total derivative of brightness E leads to

$$\frac{dE}{dt} = E_t + x_t E_x + y_t E_y = 0^1 \tag{7}$$

which is referred to as the *Brightness Change Constraint Equation* (BCCE) [15]. By substituting for x_t and y_t from eqn. (6) into eqn. (7), we obtain the brightness change constraint equation for rigid body motion [23], namely

$$E_t + \mathbf{v} \cdot \boldsymbol{\omega} + \frac{s \cdot \mathbf{t}}{Z} = 0 \tag{8}$$

where the auxiliary vectors s and $\mathbf{v}$ are defined as

$$\mathbf{s} = \begin{pmatrix} -E_x \\ -E_y \\ xE_x + yE_y \end{pmatrix} \tag{9}$$

and

$$\mathbf{v} = \begin{pmatrix} +E_y + y(xE_x + yE_y) \\ -E_x - x(xE_x + yE_y) \\ yE_x - xE_y \end{pmatrix}. \tag{10}$$

Considering that $\mathbf{s} \cdot \mathbf{r} = 0$, $\mathbf{v} \cdot \mathbf{r} = 0$ and $\mathbf{s} \cdot \mathbf{v} = 0$, these three vectors thus form an orthogonal triad. The vectors $\mathbf{s}$ and $\mathbf{v}$ represent inherent properties of the image. Also it can be shown that $\mathbf{v} = \mathbf{r} \times \mathbf{s}$. The vector $\mathbf{s}$ indicates the directions in which translation of a given magnitude will contribute maximally to the temporal brightness change of a given picture cell. The vector $\mathbf{v}$ plays a similar role for rotation.

The brightness change constraint equation is unchanged if we scale both Z and $\mathbf{t}$ by the same scale factor. We conclude that we can determine only the direction of translational velocity and the relative depth of points in the scene. This well-known ambiguity is referred to as the *scale-factor ambiguity* in motion vision.

[1] To account for smooth variations in the image brightness due to other factors such as shading, spatial and temporal illumination changes, and variations in reflectance properties, the BCCE can be extended to

$$E_t + x_t E_x + y_t E_y = m_t E + c_t$$

where in general m_t and c_t are time and position dependent [12, 25, 26]. For simplicity, we have not used this extension here but we may use it for implementation.

3 FIXATION FORMULATION

Our common visual experience suggests that fixation may play an important role in the analysis of moving objects. When we want to understand the motion of an object we do not keep our eyes and head stationary in front of the moving object. Instead our head and/or eyes follow the moving object, in order to keep the image of a point of interest stationary in the retina. There are also some formal studies that support such observations [6, 7, 8]. Consequently in this computer vision work, the *fixation* is defined as:

> Given two subsequent images, *initial 1st* and *2nd images*, and a point in the 1st image, find a sequence of two fixated images by obtaining a new image, *fixated 2nd image*, such that the image of the selected point in the new image is located at its original position as in the initial 1st image.

As shown in figure 3, we refer to this selected image point as the *fixation point*, $\mathbf{r}_\circ$, and to its corresponding point on the object as the *interest point*, $\mathbf{R}_\circ$.

3.1 Derivation of General Fixation Constraint Equation

In a sequence of two fixated images, at fixation point we should have

$$\mathbf{r}_{\circ t} = 0 \tag{11}$$

where $\mathbf{r}_{\circ t}$ is the time derivative of the fixation point vector and similar to eqn. (5) can be written as

$$\mathbf{r}_{\circ t} = \frac{\hat{\mathbf{z}} \times (\mathbf{R}_{\circ t} \times \mathbf{r}_\circ)}{\mathbf{R}_\circ \cdot \hat{\mathbf{z}}}. \tag{12}$$

$\mathbf{R}_{\circ t}$ is the time derivative of the interest point vector. Combination of equations (11) and (12) shows that for fixation we need to have

$$\hat{\mathbf{z}} \times (\mathbf{R}_{\circ t} \times \mathbf{r}_\circ) = 0. \tag{13}$$

In other words, we want to find out when $\mathbf{R}_{\circ t} \times \mathbf{r}_\circ$ is zero or parallel to $\hat{\mathbf{z}}$. For $\mathbf{R}_{\circ t} \times \mathbf{r}_\circ$ to be parallel to $\hat{\mathbf{z}}$, we should have $\mathbf{r}_\circ$ perpendicular to $\hat{\mathbf{z}}$ which is impossible with a finite field of view (FOV), so only $\mathbf{R}_{\circ t} \times \mathbf{r}_\circ = 0$ applies.

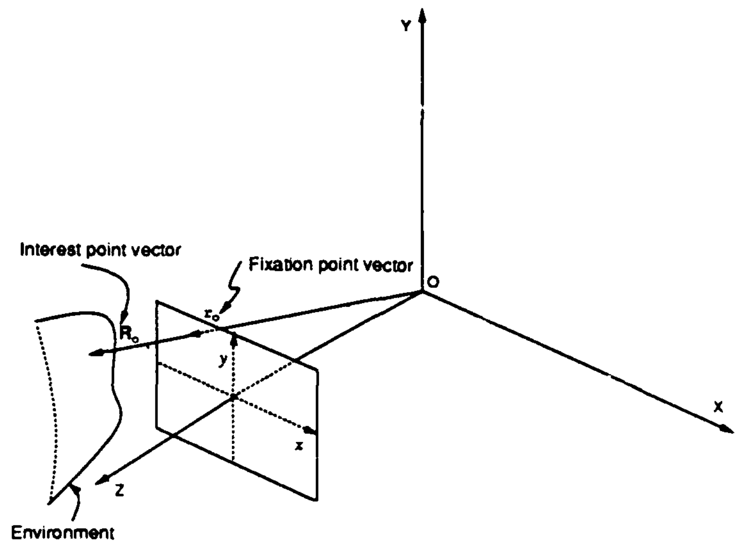

Figure 3: In *fixation* the *fixation point*, the image of the *interest point*, is kept stationary in the image plane despite the relative motion between the camera and the environment.

Conclusively, considering that $\mathbf{R}_o$ and $\mathbf{r}_o$ have the same direction, eqn. (13) is simplified as

$$\mathbf{R}_{ot} \times \mathbf{R}_o = 0 \tag{14}$$

Now substituting for $\mathbf{R}_{ot} = -\mathbf{t} - \boldsymbol{\omega} \times \mathbf{R}_o$, eqn. (4), into (14) gives

$$(\boldsymbol{\omega} \times \mathbf{R}_o) \times \mathbf{R}_o + \mathbf{t} \times \mathbf{R}_o = 0. \tag{15}$$

Expansion of eqn. (15) by using $(\mathbf{a} \times \mathbf{b}) \times \mathbf{c} = (\mathbf{c} \cdot \mathbf{a})\mathbf{b} - (\mathbf{c} \cdot \mathbf{b})\mathbf{a}$ results in

$$(\mathbf{R}_o \cdot \boldsymbol{\omega})\mathbf{R}_o - (\mathbf{R}_o \cdot \mathbf{R}_o)\boldsymbol{\omega} + \mathbf{t} \times \mathbf{R}_o = 0. \tag{16}$$

As long as the translational velocity $\mathbf{t}$ is neither zero nor parallel to the interest point vector $\mathbf{R}_o$, then any vector, including $\boldsymbol{\omega}$, can be expressed in terms of the triad of vectors $\mathbf{R}_o$, $\mathbf{t} \times \mathbf{R}_o$ and $\mathbf{t}$. So we can write $\boldsymbol{\omega}$ in its general form as

$$\boldsymbol{\omega} = \alpha \mathbf{R}_o + \beta(\mathbf{t} \times \mathbf{R}_o) + \gamma \mathbf{t} \tag{17}$$

where α, β and γ are constants to be determined. Later in this section we will consider the special cases where $\mathbf{t}$ is zero or parallel to $\mathbf{R}_\circ$ by defining ω based on another triad.

Substituting for $\boldsymbol{\omega}$ from eqn. (17) into eqn. (16) gives

$$[1 - \beta(\mathbf{R}_\circ \cdot \mathbf{R}_\circ)](\mathbf{t} \times \mathbf{R}_\circ) + \gamma(\mathbf{R}_\circ \cdot \mathbf{t})\mathbf{R}_\circ - \gamma(\mathbf{R}_\circ \cdot \mathbf{R}_\circ)\mathbf{t} = 0. \tag{18}$$

Now, we should find the constants β and γ such that eqn. (18) holds without putting any restrictions on $\mathbf{R}_\circ$ and $\mathbf{t}$. We start by finding the dot product of eqn. (18) by $\mathbf{t} \times \mathbf{R}_\circ$ that results in

$$[1 - \beta(\mathbf{R}_\circ \cdot \mathbf{R}_\circ)]||\mathbf{t} \times \mathbf{R}_\circ||^2 = 0. \tag{19}$$

Equation (19) will hold without restricting $\mathbf{R}_\circ$ and $\mathbf{t}$ if

$$\beta = \frac{1}{||\mathbf{R}_\circ||^2}. \tag{20}$$

Another possibility for satisfying eqn. (19) is to have $||\mathbf{t} \times \mathbf{R}_\circ|| = 0$ which implies that $\mathbf{t} = 0$, $\mathbf{R}_\circ = 0$ or $\mathbf{t}$ is parallel to $\mathbf{R}_\circ$. But $\mathbf{R}_\circ$ cannot be zero and also we assumed that here $\mathbf{t}$ is neither zero nor parallel to $\mathbf{R}$o. As a result, $||\mathbf{t} \times \mathbf{R}_\circ||$ cannot be zero. Similarly the dot product of eqn. (18) by $\mathbf{t}$ gives

$$\gamma(\mathbf{R}_\circ \cdot \mathbf{t})(\mathbf{R}_\circ \cdot \mathbf{t}) - \gamma(\mathbf{R}_\circ \cdot \mathbf{R}_\circ)(\mathbf{t} \cdot \mathbf{t}) = 0. \tag{21}$$

Knowing that $(\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{c} \times \mathbf{d}) = (\mathbf{c} \cdot \mathbf{a})(\mathbf{b} \cdot \mathbf{d}) - (\mathbf{d} \cdot \mathbf{a})(\mathbf{b} \cdot \mathbf{c})$, eqn. (21) can be simplified as

$$\gamma||\mathbf{t} \times \mathbf{R}_\circ||^2 = 0. \tag{22}$$

We discussed that $||\mathbf{t} \times \mathbf{R}_\circ||$ cannot be zero here, so eqn. (22) is satisfied only if γ is zero

$$\gamma = 0. \tag{23}$$

Substituting for β from eqn. (20) and γ from eqn. (23) into eqn. (17) gives

$$\boldsymbol{\omega} = \alpha\mathbf{R}_\circ + \frac{1}{||\mathbf{R}_\circ||^2}(\mathbf{t} \times \mathbf{R}_\circ) \tag{24}$$

where α is still unknown. This means that the component of the rotational velocity along $\mathbf{R}_\circ$ cannot be determined by the *fixation formulation*. Physically this makes sense because the rotational velocity along $\mathbf{R}_\circ$, denoted by

$\omega_{\mathbf{R}_\circ}$, does not move the fixation point. This observation leads us to find $\omega_{\mathbf{R}_\circ}$ in a separate step before using the fixation formulation results. Derivation of $\omega_{\mathbf{R}_\circ}$ will be shown in section 5. Finally the *general fixation constraint equation* (GFCE) is written as

$$\omega = \omega_{\mathbf{R}_\circ}\hat{\mathbf{R}}_\circ + \frac{1}{\|\mathbf{R}_\circ\|}(\mathbf{t} \times \hat{\mathbf{R}}_\circ) \tag{25}$$

where $\mathbf{t}$ is the translational velocity and $\hat{\mathbf{R}}_\circ = \hat{\mathbf{r}}_\circ$ is the unit vector along the position vector of an arbitrary fixation point, a point in the image chosen for fixation.

3.2 Derivation of Special Fixation Constraint Equation

When the translational velocity $\mathbf{t}$ is zero or parallel to the interest point vector $\mathbf{R}_\circ$, eqn. (16) is simplified as

$$(\mathbf{R}_\circ \cdot \omega)\mathbf{R}_\circ - (\mathbf{R}_\circ \cdot \mathbf{R}_\circ)\omega = 0. \tag{26}$$

We define $\boldsymbol{\omega}$ based on the triad consisting of vectors $\mathbf{R}_\circ$, $\hat{\mathbf{x}}$, and $\hat{\mathbf{x}} \times \mathbf{R}_\circ$ as

$$\omega = l\mathbf{R}_\circ + m(\hat{\mathbf{x}} \times \mathbf{R}_\circ) + n\hat{\mathbf{x}} \tag{27}$$

where l, m, and n are constants to be determined. Here we assume that $\mathbf{R}_\circ$ is not parallel to $\hat{\mathbf{x}}$. This is a reasonable assumption because otherwise we should at least have a field of view of 180° to be able to choose an awkward point of interest along the x-axis, which results in a fixation point at infinite distance from the principal point and near the border of an infinite image plane.

Substituting for $\boldsymbol{\omega}$ from eqn. (27) into eqn. (26) gives

$$[m\mathbf{R}_\circ \cdot (\hat{\mathbf{x}} \times \mathbf{R}_\circ) + n(\mathbf{R}_\circ \cdot \hat{\mathbf{x}})]\mathbf{R}_\circ - m(\mathbf{R}_\circ \cdot \mathbf{R}_\circ)(\hat{\mathbf{x}} \times \mathbf{R}_\circ) - n(\mathbf{R}_\circ \cdot \mathbf{R}_\circ)\hat{\mathbf{x}} = 0. \tag{28}$$

The dot product of eqn. (28) with $(\hat{\mathbf{x}} \times \mathbf{R}_\circ)$ results in

$$-m(\mathbf{R}_\circ \cdot \mathbf{R}_\circ)\|\hat{\mathbf{x}} \times \mathbf{R}_\circ\|^2 = 0. \tag{29}$$

Considering that $\mathbf{R}_\circ$ cannot be either zero or parallel to $\hat{\mathbf{x}}$, eqn. (29) is satisfied only if m is zero

$$m = 0. \tag{30}$$

Substituting for m into eqn. (28) and finding its dot product by $\hat{\mathbf{x}}$ results in

$$n(\mathbf{R}_\circ \cdot \hat{\mathbf{x}})(\mathbf{R}_\circ \cdot \hat{\mathbf{x}}) - n(\mathbf{R}_\circ \cdot \mathbf{R}_\circ)(\hat{\mathbf{x}} \cdot \hat{\mathbf{x}}) = 0. \tag{31}$$

Using $(\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{c} \times \mathbf{d}) = (\mathbf{c} \cdot \mathbf{a})(\mathbf{b} \cdot \mathbf{d}) - (\mathbf{d} \cdot \mathbf{a})(\mathbf{b} \cdot \mathbf{c})$, eqn. (31) can be written as

$$n\|\hat{\mathbf{x}} \times \mathbf{R}_\circ\|^2 = 0. \tag{32}$$

Again $\mathbf{R}_\circ$ cannot be either zero or parallel to $\hat{\mathbf{x}}$. As a result, eqn. (32) will hold for arbitrary $\mathbf{R}_\circ$ if $n = 0$. Substituting for n and m into eqn. (27) gives

$$\omega = l\mathbf{R}_\circ \tag{33}$$

where l is still unknown. We can substitute $\omega_{\mathbf{R}_\circ}\hat{\mathbf{R}}_\circ$ for $l\mathbf{R}_\circ$. It will be shown later how $\omega_{\mathbf{R}_\circ}$ is found separately. As a result for the special cases we obtain the *special fixation constraint equation* (SFCE) as

$$\omega = \omega_{\mathbf{R}_\circ}\hat{\mathbf{R}}_\circ \tag{34}$$

which means that when the translational velocity $\mathbf{t}$ is zero or parallel to $\mathbf{R}_\circ$ then the corresponding rotational velocity may only have a component along $\mathbf{R}_\circ$.

Our method for deriving the SFCE, eqn. (34), is not different from what we did for deriving the GFCE, eqn. (25). In fact, eqn. (34) is a special case of eqn. (25). But we could not directly derive eqn. (34) from eqn. (25) because eqn. (25) was derived based on the assumption that $\mathbf{t}$ is neither zero nor parallel to $\mathbf{R}_\circ$. As a result, for implementation it is enough to use the GFCE, eqn. (25), without knowing whether the present condition is the special case or not.

4 SOLVING THE GENERAL DIRECT MOTION VISION PROBLEM

Here we assume that we are given a sequence of fixated images. In other words we have made the fixation point stationary in our image sequence. This can be done by finding the *fixation velocity*, the apparant velocity at the fixation point in the 1st image, as given in section 5. Then the shifting method explained in section 6 can be used for generating a new image, *fixated 2nd image*, in which the fixation point is located at the same position as in the initial 1st image.

We start by studying the general case where the translational velocity $\mathbf{t}$ is neither zero nor parallel to the interest point vector $\mathbf{R}_\circ$. Then we will consider the special cases of $\mathbf{t}$ separately.

Substituting for $\boldsymbol{\omega}$ from the general fixation constraint equation (25) into the brightness-change constraint equation (8) gives

$$E_{\mathbf{t}} + \omega_{\mathbf{R}_\circ}\mathbf{v}\cdot\hat{\mathbf{R}}_\circ + \frac{1}{\|\mathbf{R}_\circ\|}[\mathbf{v}\cdot(\mathbf{t}\times\hat{\mathbf{R}}_\circ)] + \frac{1}{Z}(\mathbf{s}\cdot\mathbf{t}) = 0. \tag{35}$$

Knowing that $\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c}) = (\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}$ and doing some manipulations on eqn. (35) results in

$$E_t' + [\frac{1}{Z}\mathbf{s} - \frac{1}{\|\mathbf{R}_\circ\|}(\mathbf{v}\times\hat{\mathbf{R}}_\circ)]\cdot\mathbf{t} = 0 \tag{36}$$

where E_t' is a notation for $E_t + \omega_{\mathbf{R}_\circ}\mathbf{v}\cdot\hat{\mathbf{R}}_\circ$ which can be computed at any pixel. In general, eqn. (36) can be solved numerically for $\mathbf{t}$ and Z using images of any size and with any field of view (FOV). In the following, a closed form solution is presented for the case that a small patch around the fixation point is used or the field of view is small and the whole image is used.

We know that at the fixation point $\mathbf{v}\times\hat{\mathbf{R}}_\circ = \mathbf{v}_\circ\times\hat{\mathbf{R}}_\circ = 0$. For a small field of view, the product of $\mathbf{v}\times\hat{\mathbf{R}}_\circ$ will be negligible. Even for an image with a large field of view this is still true for the image area near the fixation point. As a result, for these cases, eqn. (36) can be written as

$$E_t' + \frac{1}{Z}(\mathbf{s}\cdot\mathbf{t}) \approx 0 \tag{37}$$

which can be solved similar to the pure translation case [17, 24, 16]. Questions such as "How large can the FOV or fixation patch be to guarantee the validity of this approximation?" can be answered by implementation results.

Considering that the depth range is finite, we can solve eqn. (37) by minimizing

$$\iint Z^2 dx dy = \iint (\frac{\mathbf{s}\cdot\mathbf{t}}{E_t'})^2 dx dy = J \tag{38}$$

with respect to $\mathbf{t}$. In other words, we are looking for the true motion $\mathbf{t}$ which minimizes the sum of squares of the depth estimates over the image of a scene with a finite depth range. In order to avoid the trivial solution $\mathbf{t} = 0$, a constraint such as $\|\mathbf{t}\| = 1$ is put on this minimization problem. This is a

valid constraint on $\mathbf{t}$ because due to the *scale factor ambiguity* we can only find the direction of $\mathbf{t}$. This constraint on $\mathbf{t}$ can be written as

$$\mathbf{t}^T\mathbf{t} = 1. \tag{39}$$

Moreover we can rewrite J as

$$J = \mathbf{t}^T M \mathbf{t} \tag{40}$$

where M is a fully computable 3×3 matrix

$$M = \iint (\frac{1}{E_t'})^2 \mathbf{s}\mathbf{s}^T dxdy. \tag{41}$$

Minimizing J in eqn. (40) under the constraint eqn. (39) is an ordinary calculus constrained minimization problem which can be solved by minimizing

$$I(\mathbf{t},\ \lambda) = \mathbf{t}^T M \mathbf{t} + \lambda(1 - \mathbf{t}^T\mathbf{t}) \tag{42}$$

with respect to $\mathbf{t}$ and the Lagrange multiplier λ. Then we will have

$$\frac{\partial I}{\partial t} = 2Mt - 2\lambda\mathbf{t} = 0 \tag{43}$$

which is simplified as

$$M\mathbf{t} = \lambda\mathbf{t}. \tag{44}$$

Equation (44) is an eigenvalue problem where λ is an eigenvalue of the known matrix M and $\mathbf{t}$ is the corresponding eigenvector. Substituting $M\mathbf{t}$ from eqn. (44) into eqn. (42) gives $I = \lambda$ which implies that under the given constraint $\mathbf{t}^T M\mathbf{t}$ is minimized when the smallest of three eigenvalues is used for calculating the eigenvector $\mathbf{t}$.

It is concluded that the fixation method can be used for solving the motion vision problem in its general case. The translational velocity $\mathbf{t}$ can be calculated from eqn. (44) by using the smallest eigenvalue. Then we can use eqn. (36) for finding the environment depth

$$Z = -\frac{1}{E_t' - \frac{(\mathbf{v}\times\hat{\mathbf{R}}_o)\cdot\mathbf{t}}{\|\mathbf{R}_o\|}}(\mathbf{s}\cdot\mathbf{t}) \approx -\frac{1}{E_t'}(\mathbf{s}\cdot\mathbf{t}) \tag{45}$$

and finally eqn. (25) gives the rotational velocity $\boldsymbol{\omega}$

$$\boldsymbol{\omega} = \omega_{\mathbf{R}_\circ}\hat{\mathbf{R}}_\circ + \frac{1}{\|\mathbf{R}_\circ\|}(\mathbf{t} \times \hat{\mathbf{R}}_\circ). \tag{46}$$

The total rotational velocity of the vehicle with respect to the environment is obtained by adding $\boldsymbol{\omega}$ to the *equivalent rotational velocity* $\boldsymbol{\Omega}$ given in section 6. It can be seen that for the general case, the *fixation formulation* lets us find the shape and motion parameters based on an arbitrary choice of interest point $\mathbf{R}_\circ$.

4.1 Special Cases: t Is Zero or Parallel to $\mathbf{R}_\circ$

When the translational velocity $\mathbf{t}$ is zero, we showed that the rotational velocity $\boldsymbol{\omega}$ has only a component along $\mathbf{R}_\circ$. In this case we basically cannot obtain any estimation for the depth Z but there are methods for finding the rotational velocity $\boldsymbol{\omega}$ [17]. For the other special case where $\mathbf{t}$ is parallel to $\mathbf{R}_\circ$, we substitute for $\boldsymbol{\omega}$ from eqn. (34) into the BCCE eqn. (8) to obtain

$$E'_t + \frac{1}{Z}(\mathbf{s} \cdot \mathbf{t}) = 0 \tag{47}$$

where E'_t is a notation for the computable value of $E_t + \omega_{\mathbf{R}_\circ}\mathbf{v} \cdot \hat{\mathbf{R}}_\circ$. Because no approximation is involved in deriving eqn. (47), an exact closed form solution exists for $\mathbf{t}$ and Z without any restriction on the field of view or the image size. This exact solution for finding $\mathbf{t}$ and Z is the same as the solution given in the general case, starting from eqn. (38).

5 COMPUTING THE FIXATION VELOCITY AND $\omega_{\mathbf{R}_\circ}$

The fixation formulation is based on the assumption that the fixation point remains stationary in a sequence of fixated images. We use the term *fixation velocity* to refer to the apparent velocity at the fixation point in the initial 1st image. We also represent x and y components of the fixation velocity by $u_\circ$ and $v_\circ$ respectively. The basic fixation requirement, a sequence of two fixated images in which $\mathbf{r}_{\circ t} = 0$, can be satisfied by finding $u_\circ$ and $v_\circ$, and then using these components for obtaining a new image, *fixated 2nd image*. The *shifting method* for obtaining the fixated 2nd image is explained in the next section.

We also saw that the component of the rotational velocity along $R_\circ$, $\omega_{\mathbf{R}_\circ}$, cannot be calculated from the fixation formulation because this component does not move the fixation point. Here, we will introduce algorithms which can be used for finding both $\omega_{\mathbf{R}_\circ}$ and the components of the fixation velocity.

If we assume that depth is approximately constant on a small patch around the fixation point, the *fixation patch*, then $u_\circ$ and $v_\circ$ will be approximately constant on this patch. Possible sensitivity of this assumption to special cases such as slanted surfaces can be checked by implementation. Moreover the motion field velocity due to the component of the rotational velocity of the camera with respect to the environment along $\mathbf{R}_\circ$ is given by $-(\boldsymbol{\omega}_{\mathbf{R}_\circ} \times \mathbf{r}) = -\omega_{\mathbf{R}_\circ}(\hat{\mathbf{R}}_\circ \times \mathbf{r}) = -\frac{\omega_{\mathbf{R}_\circ}}{\|\mathbf{r}_\circ\|}(\mathbf{r}_\circ \times \mathbf{r})$ because $\hat{\mathbf{R}}_\circ = \hat{\mathbf{r}}_\circ$ is the unit vector along $\mathbf{r}_\circ$. Knowing that $\mathbf{r}_\circ = (x_\circ\ y_\circ\ 1)^T$ and $\mathbf{r} = (x\ y\ 1)^T$, the components of the total motion field velocity along the x and y axes, due to fixation velocity and $\omega_{\mathbf{R}_\circ}$, are given by

$$\left\{ \begin{array}{lllll} x_t & = & u_\circ - \frac{\omega_{\mathbf{R}_\circ}}{\|\mathbf{r}_\circ\|}\hat{\mathbf{x}} \cdot (\mathbf{r}_\circ \times \mathbf{r}) & = & u_\circ - \bar{\omega}_{\mathbf{R}_\circ}(y_\circ - y) \\ y_t & = & v_\circ - \frac{\omega_{\mathbf{R}_\circ}}{\|\mathbf{r}_\circ\|}\hat{\mathbf{y}} \cdot (\mathbf{r}_\circ \times \mathbf{r}) & = & v_\circ - \bar{\omega}_{\mathbf{R}_\circ}(x - x_\circ) \end{array} \right. \tag{48}$$

where $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ are the unit vectors along the x and y axes and $\bar{\omega}_{\mathbf{R}_\circ}$ is a notation for $\frac{\omega_{\mathbf{R}_\circ}}{\|\mathbf{r}_\circ\|}$. Substituting for x_t and y_t from the above equations into the BCCE, eqn. (7), gives

$$[u_\circ - \bar{\omega}_{\mathbf{R}_\circ}(y_\circ - y)]E_x + [v_\circ - \bar{\omega}_{\mathbf{R}_\circ}(x - x_\circ)]E_y + E_t = 0. \tag{49}$$

Due to noise, eqn. (49) does not necessarily hold for any $\mathbf{r}$ so we try to find $u_\circ, v_\circ$ and $\bar{\omega}_{\mathbf{R}_\circ}$ by minimizing the sum of squares of errors over the fixation patch, denoted by p. In other words we want to minimize

$$\iint_p [(u_\circ - \bar{\omega}_{\mathbf{R}_\circ}(y_\circ - y))E_x + (v_\circ - \bar{\omega}_{\mathbf{R}_\circ}(x - x_\circ))E_y + E_t]^2 dx\, dy \tag{50}$$

with respect to $u_\circ$, $v_\circ$ and $\bar{\omega}_{\mathbf{R}_\circ}$ which results in a system of three linear equations which can be solved for the three unknowns

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{pmatrix} u_\circ \\ v_\circ \\ \bar{\omega}_{\mathbf{R}_\circ} \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}. \tag{51}$$

Matrix $\mathbf{A}$ is symmetric and its elements are given by

$$\left\{\begin{array}{rcrcl} a_{12} & = & a_{21} & = & \iint_p E_x E_y dx\, dy \\ a_{13} & = & a_{31} & = & -\iint_p E_x[E_x(y_\circ - y) + E_y(x - x_\circ)]dx\, dy \\ a_{23} & = & a_{32} & = & -\iint_p E_y[E_x(y_\circ - y) + E_y(x - x_\circ)]dx\, dy \\ & & a_{11} & = & \iint_p E_x^2 dx\, dy \\ & & a_{22} & = & \iint_p E_y^2 dx\, dy \\ & & a_{33} & = & \iint_p [E_x(y_\circ - y) + E_y(x - x_\circ)]^2 dx\, dy \end{array}\right. \tag{52}$$

and the elements of vector C are as follows:

$$\left\{\begin{array}{rcl} c_1 & = & -\iint_p E_t E_x dx\, dy \\ c_2 & = & -\iint_p E_t E_y dx\, dy \\ c_3 & = & \iint_p E_t[E_x(y_\circ - y) + E_y(x - x_\circ)]dx\, dy. \end{array}\right. \tag{53}$$

Considering that the fixation point coordinates $x_\circ$ and $y_\circ$ are known, then the sets of equations (52) and (53) show that the elements of matrix A and vector C can be calculated easily.

In the special case where the fixation point is at the principal point, $x_\circ = y_\circ = 0$, the a_{ij} are simplified as:

$$\left\{\begin{array}{rcrcl} a_{12} & = & a_{21} & = & \iint_p E_x E_y dx\, dy \\ a_{13} & = & a_{31} & = & \iint_p E_x(yE_x - xE_y)dx\, dy \\ a_{23} & = & a_{32} & = & \iint_p E_y(yE_x - xE_y)dx\, dy \\ & & a_{11} & = & \iint_p E_x^2 dx\, dy \\ & & a_{22} & = & \iint_p E_y^2 dx\, dy \\ & & a_{33} & = & \iint_p (yE_x - xE_y)^2 dx\, dy \end{array}\right. \tag{54}$$

and c_i are given as follows:

$$\left\{\begin{array}{rcl} c_1 & = & -\iint_p E_t E_x dx\, dy \\ c_2 & = & -\iint_p E_t E_y dx\, dy \\ c_3 & = & -\iint_p E_t(yE_x - xE_y)dx\, dy. \end{array}\right. \tag{55}$$

After finding $\bar{\omega}_{\mathbf{R}_\circ}$, we can easily calculate $\omega_{\mathbf{R}_\circ}$ as

$$\omega_{\mathbf{R}_\circ} = \bar{\omega}_{\mathbf{R}_\circ}\sqrt{x_\circ^2 + y_\circ^2 + 1}. \tag{56}$$

When the fixation point is at the principal point, $\omega_{\mathbf{R}_\circ}$ is exactly the same as $\bar{\omega}_{\mathbf{R}_\circ}$.

6 OBTAINING A FIXATED 2ND IMAGE

The fixation method assumes that a sequence of two images are available in which the fixation point is kept stationary. Referring to figure 1, we are given two original images. The 1st original image is used directly but we need to find a *fixated 2nd image.*

Physical rotation of the camera with respect to the vehicle is a hardware solution to this problem which is basically a tracking problem. Considering that in general the interest point has a relative motion with respect to the vehicle, the fixated 2nd image cannot be obtained in one step. As a result, a feedback loop is required for the camera rotation system to compensate for the errors resulted from the new position of the fixation point. This hardware approach is avoided not only because of the errors involved but also because of the concern about the real time applications.

In the following we will show how a fixated 2nd image can be obtained by applying a compensating rotation to the initial 2nd image through software. It is assumed that the fixation velocity has been already computed, eqn. (51). We introduce an *equivalent rotational velocity* $\mathbf{\Omega} = (\Omega_x, \Omega_y, \Omega_z)$ which could result in the same fixation velocity (u_o, v_o) at the fixation point (x_o, y_o). According to eqn. (6), the components of $\mathbf{\Omega}$ must satisfy the following set of equations:

$$\begin{cases} u_o = x_o y_o \Omega_x - (x_o^2 + 1)\Omega_y + y_o \Omega_z \\ v_o = (y_o^2 + 1)\Omega_x - x_o y_o \Omega_y - x_o \Omega_z. \end{cases} \tag{57}$$

Among infinite number of $\mathbf{\Omega}$ that satisfy the system of equations (57), we choose the only one that does not result in any rotational velocity along the fixation point vector $\mathbf{r}_o$. Mathematically this means that $\mathbf{\Omega} \cdot r_o = 0$ which results in the following constraint on the components of $\mathbf{\Omega}$

$$x_o \Omega_x + y_o \Omega_y + \Omega_z = 0. \tag{58}$$

Given the fixation velocity (u_o, v_o) and the fixation point coordinates x_o and y_o, the equivalent rotational velocity $\mathbf{\Omega}$ is obtained by solving the combination of three linear equations in (57) and (58). In the case that the fixation point is at the principal point, $x_o = y_o = 0$, the equivalent rotational velocity is

$$\mathbf{\Omega} = (v_o, -u_o, 0). \tag{59}$$

Let's define the *rotational fixation velocity* as

$$\mathbf{\Omega}_o = (\Omega_{x_o}, \Omega_{y_o}, \Omega_{z_o}) = -\mathbf{\Omega}. \tag{60}$$

In other words $\mathbf{\Omega}_o$ is equal to but in opposite direction of the equivalent rotational velocity $\mathbf{\Omega}$ given by equations (57) and (58). The 2nd fixated image can be obtained by applying $\mathbf{\Omega}_o$ to the initial 2nd image. Considering eqn. (57), the following set of equations must be satisfied in the shifting process of the initial 2nd image

$$\begin{cases} u = xy\Omega_{x_o} - (x^2+1)\Omega_{y_o} + y\Omega_{z_o} \\ v = (y^2+1)\Omega_{x_o} - xy\Omega_{y_o} - x\Omega_{z_o}. \end{cases} \tag{61}$$

Here Ω_{x_o}, Ω_{y_o} and Ω_{z_o} are known, as a result the *shifting vector* (u, v) can be obtained for every pixel of the initial 2nd image. The brightness at pixel (x, y) of the fixated 2nd image is obtained by finding the brightness at the corresponding original point $(x - Tu, y - Tv)$ in the initial 2nd image. T is the time interval between two initial images. In general a computed original point is not located at the center of a pixel in the initial 2nd image. As a result, its brightness cannot be read directly from the image file and should be computed by averaging, bilinear interpolation or bicubic interpolation of the brightnesses at its neighboring pixels.

7 CONCLUSIONS AND FUTURE WORK

The algorithms and formulations presented in this *fixation method* show how to solve the motion vision problem directly for arbitrary motion with respect to an arbitrary rigid scene. In contrast to previous work done in the area of motion vision, this solution is general and does not put any severe restriction on the motion or the shape of environment. More importantly the fixation method uses neither *optical flow* nor *feature correspondence*; instead direct image information such as temporal and spatial brightness gradients are used. There is no restriction on choosing the *fixation point*. However using the principal point as the fixation point makes the equations more concise and the calculations easier.

Implementation of this fixation method, which will be our next work, is essential for supporting the feasibility of the scheme. Referring to fig. 1, we can implement the fixation method in the following steps.

STEP 1: Finding the ***fixation velocity*** components (u_o, v_o) and component of rotational velocity along $\mathbf{R}_o$, $\omega_{\mathbf{R}_o}$, by applying the system of eqn. (51) to the direct image information from two *initial images.*

STEP 2: Knowing the fixation velocity components, u_o and v_o, the ***fixated 2nd image*** is obtained by the ***shifting method*** explained in section 6.

STEP 3: Using the general fixation constraint equation (25), original 1st image, and the fixated 2nd image, the method presented in section 4 can be used for recovering the depth, the translational velocity and the rotational velocity. Note that we had to derive the special fixation constraint equation (34) separately, but for implementation it is enough to use just the general fixation constraint equation (25) because eqn. (34) is a special case of eqn. (25). As a result, the general algorithms can be used for recovering the motion and depth, without knowing in advance whether the present condition is a special case or not.

STEP 4: The *total rotational velocity* $\boldsymbol{\omega}_{tot}$ is simply obtained by adding the *equivalent rotational velocity* $\boldsymbol{\Omega}$, from equations (57) and (58), to the *rotational velocity* $\boldsymbol{\omega}$ from eqn. (46).

8 ACKNOWLEDGEMENT

I would like to express my sincere gratitude to Prof. Berthold K. P. Horn who patiently went through several versions of this paper and provided me with many valuable ideas. Also many thanks to Prof. Ellen C. Hildreth for her crucial, constructive, and very useful comments on this work.

References

[1] Adiv, G. "Determining 3-D Motion and Structure from Optical Flow Generated by Several Moving Objects," *IEEE Transaction on Pattern Analysis and Machine Intelligence,* Vol. 7, No 4, July 1985.

[2] Aloimonos, J. & Basu, A. "Determining the Translation of a Rigidly Moving Surface, without Correspondence," *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition,* Miami, FL, June 1986.

[3] Aloimonos, J. & Tsakiris, D. P. "On the Mathematics of Visual Tracking," *Computer Vision Laboratory, University of Maryland, Maryland,* Rep. CAR-TR-390, September 1988.

[4] Aloimonos, J. & Shulman, D. *Integration of Visual Modules, An Extension of the Marr Paradigm*, Boston, Academic Press, 1989.

[5] Ballard, D. H. & Kimball, O. A. "Rigid Body Motion from Depth and Optical Flow," *Computer Vision, Graphics, and Image Processing,* Vol. 22, No 1, April 1983.

[6] Bahill, A.T. & LaRitz, T. "Why Can't Batters Keep their Eyes on the Ball?," *American Scientist,* Vol. 72, pp. 219-253, 1984.

[7] Bahill, A.T. & McDonald, J.D. "Model Emulates Human Smooth Pursuit System Producing Zero-Latency Target Tracking," *Biol. Cybern.,* Vol. 48, pp. 213-222, 1983.

[8] Bandopadhay, A. "A Computational Study of Rigid Motion Perception," *Ph.D. Thesis, Department of Computer Science, University of Rochester,* 1986.

[9] Bandopadhay, A., Chandra, B. & Ballard, D.H. "Active Navigation: Tracking an Environmental Point Considered Beneficial." *Proc. of IEEE Workshop on Motion: Representation and Analysis,* Kiawash Island, pp. 23-29, May 7-9, 1986.

[10] Barron J. "A Survey of Approaches for Determining Optical Flow, Environmental Layout and Egomotion," *University of Toronto, Ontario, Canada,* Rep. RBCV-TR-84-5, November 1984.

[11] Bruss, A. R. & Horn, B. K. P. "Passive Navigation," *Computer Vision, Graphics, and Image Processing,* Vol. 21, No 1, pp. 3-20, January 1983.

[12] Gennert, M. A. & Negahdaripour, S. "Relaxing the Brightness Constancy Assumption in Computing Optical Flow," *MIT, AI Lab.,* Cambridge, MA, AI Memo 975, June 1987.

[13] Hallert, B. *Photogrammetry,* McGraw Hill, New York, NY, 1960.

[14] Hildreth, E. C. *The Measurement of Visual Motion,* Cambridge, MA, MIT Press, 1983.

[15] Horn, B. K. P. & Schunck, B. G. "Determining Optical Flow," *Artificial Intelligence,* Vol. 17, pp. 185-203, 1981.

[16] Horn, B. K. P. & Weldon, E. J. Jr. "Computationally Efficient Methods of Recovering Translational Motion," *International Conference of Computer Vision,* London, England, June 1987.

[17] Horn, B. K. P. & Weldon, E. J. Jr. "Direct Methods for Recovering Motion," *International Journal of Computer Vision.* Vol. 2, pp. 51-76, 1988.

[18] Longuet-Higgins, H. C. & Prazdny, K. "The Interpretation of a Moving Retinal Image," *Proceeding of the Royal Society of London,* B 208 pp. 385-397, 1980.

[19] Longuet-Higgins, H. C. "A Computer Algorithm for Reconstructing a Scene from Two Projections," *Nature,* Vol. 293, 1981.

[20] Moffit, F. & Mikhail, E. M. *Photogrammetry,* 3-rd edition, Harper & Row, New York, NY, 1980.

[21] Negahdaripour, S. & Horn, B. K. P. "Direct Passive Navigation," *MIT, AI Lab.,* Cambridge, MA, AI Memo 821, February 1985.

[22] Negahdaripour, S. "Direct Methods for Structure from Motion," *Ph.D. Thesis, Mechanical Engineering, MIT.* September 1986.

[23] Negahdaripour, S. & Horn, B. K. P. "Direct Passive Navigation," *IEEE Transaction on Pattern Analysis and Machine Intelligence,* Vol. 9, No 1, pp. 168-176, January 1987.

[24] Negahdaripour, S. & Horn, B. K. P. "Using Depth-Is-Positive Constraint to Recover Translational Motion," *IEEE Workshop on Computer Vision.* Miami, FL, 1987.

[25] Negahdaripour, S., Shokrollahi, A. & Gennert, M. "Relaxing the Brightness Constancy Assumption in Computing Optical Flow," *Proc. of International Conf. on Image Processing,* Sigapore, pp. 806-810, September 1989.

[26] Negahdaripour, S., & Shokrollahi, A. "Robust Motion Recovery: The Two-Plane Method," *Proc. of Second Int. Conference on Intelligent Autonomous Systems,* Amsterdam, December 1989.

[27] Prazdny, K. "Motion and Structure from Optical Flow," *Proceedings of the Sixth International Joint Conference on Artificial Intelligence,* Tokyo, Japan, August 1979.

[28] Sadini, G., Tagliasco, V. & Tsitarelli, M. "Analysis of Object Motion and Camera Motion in Real Scenes," *Proc. of IEEE Int. Conf. on Robotics and Automation,* San Fransisco, CA, pp. 627-633, April 7-10, 1986.

[29] Sadini, G. & Tistarelli, M. "Active Tracking Strategy for Monocular Depth Inference Over Multiple Frames," *Pattern Analysis and Machine Intelligence,* vol. 12, No. 1, pp. 13-27, January 1990.

[30] Ullman, S. *The Interpretation of Visual Motion*, Cambridge, MA, MIT Press, 1979.

[31] Ullman, S. "The Effect of Similarity between Line Segments on the Correspondence Strength in Apparent Motion," *Perception*, Vol. 9, pp. 617-626, 1980.

[32] Waxman, A. M. & Wohn, K. "Image Flow Theory: A Frame Work for 3-D Inference from Time-Varing Imagery," *Advances in Computer Vision, Volume I*, ed. C. Brown, Lawrence Erlbaum Assoc., 1988.